Color Bar Coding the *BRCA1* Gene on Combed DNA: A Useful Strategy for Detecting Large Gene Rearrangements

Sophie Gad,[1] Alain Aurias,[2] Nadine Puget,[3] Aline Mairal,[2] Catherine Schurra,[4] Marco Montagna,[5] Sabine Pages,[1] Virginie Caux,[1] Sylvie Mazoyer,[3] Aaron Bensimon,[4] and Dominique Stoppa-Lyonnet[1]*

[1]Service de Génétique Oncologique, Institut Curie, Paris, France
[2]INSERM U509 Pathologie Moléculaire des Cancers, Institut Curie, Paris, France
[3]International Agency for Research on Cancer and Laboratoire de Génétique, Unité Mixte de Recherche 5641 CNRS, Université Claude Bernard, Lyon, France
[4]Laboratoire de Biophysique de l'ADN, Institut Pasteur, Paris, France
[5]Department of Oncology and Surgical Sciences, Oncology Section, Interuniversity Center for Research on Cancer, University of Padova, Padova, Italy

Genetic linkage data have shown that alterations of the *BRCA1* gene are responsible for the majority of hereditary breast and ovarian cancers. *BRCA1* germline mutations, however, are found less frequently than expected. Mutation detection strategies, which are generally based on the polymerase chain reaction, therefore focus on point and small gene alterations. These approaches do not allow for the detection of large gene rearrangements, which also can be involved in *BRCA1* alterations. Indeed, a few of them, spread over the entire *BRCA1* gene, have been detected recently by Southern blotting or transcript analysis. We have developed an alternative strategy allowing a panoramic view of the *BRCA1* gene, based on dynamic molecular combing and the design of a full four-color bar code of the *BRCA1* region. The strategy was tested with the study of four large *BRCA1* rearrangements previously reported. In addition, when screening a series of 10 breast and ovarian cancer families negatively tested for point mutation in *BRCA1/2*, we found an unreported 17-kb *BRCA1* duplication encompassing exons 3 to 8. The detection of rearrangements as small as 2 to 6 kb with respect to the normal size of the studied fragment is achieved when the *BRCA1* region is divided into 10 fragments. In addition, as the *BRCA1* bar code is a morphologic approach, the direct observation of complex and likely underreported rearrangements, such as inversions and insertions, becomes possible.

INTRODUCTION

Genetic predisposition associated with a high risk of breast and ovarian cancer is one of the most common pathologic genetic conditions, with a carrier frequency estimated at 1/500. Alterations of either the *BRCA1* or *BRCA2* gene, transmitted in an autosomal dominant mode, account for the majority of familial breast and ovarian cancer cases (Ford et al., 1998). Indeed, linkage analyses performed on a set of 94 families with at least four cases of breast cancer and at least one ovarian cancer case have shown that *BRCA1* and *BRCA2* are involved in 95% of families, *BRCA1* being altered in 80% of such cases (Ford et al., 1998). Although numerous efforts in screening for *BRCA1* have been made, only two-thirds of expected mutations are detected at present (Gayther and Ponder, 1997; Stoppa-Lyonnet et al., 1997; Ford et al., 1998). Most reported approaches (single-strand conformation polymorphism, heteroduplex analysis, denaturing gradient gel electrophoresis, protein truncation test, direct sequencing) are based on polymerase chain reaction (PCR) techniques and therefore focus on point and small mutations (Gayther and Ponder, 1997).

The low *BRCA1* mutation detection rate may be due in part to the existence of large rearrangements not detected by such approaches. Indeed, PCR-based methods can miss an altered allele when it is hidden by the persisting normal allele. Supporting this hypothesis, 15 different large rearrangements have been detected recently by Southern blotting or analysis of *BRCA1* lymphocyte transcripts (Payne and King, 1997; Petrij-Bosch et al., 1997; Puget et al., 1997, 1999a, b; Swensen et al., 1997; Carson et al., 1999; Montagna et al., 1999; Rohlfs et al., 1999, 2000). The characterized rearrangements

Supported by: Institut Curie "Programme Incitatif et Coopératif: Génétique et Biologie des Cancers du Sein"; MENRT; Fondazione Cassa di Risparmio di Padova e Rovigo.

*Correspondence to Dominique Stoppa-Lyonnet, Service de Génétique Oncologique-Institut Curie, Section Médicale, 26, rue d'Ulm, F-75248 Paris Cedex 05, France.
E-mail: dominique.lyonnet@curie.net

Received 1 August 2000; Accepted 31 October 2000
Published online 8 March 2001.

are due to unequal recombination events between *Alu* sequences that cover 41.5% of *BRCA1* introns (Smith et al., 1996). First estimations of the frequency of large rearrangements in the *BRCA1* mutation spectrum range from 15% to 36%, the highest value reflecting founder effects in the Dutch population (Petrij-Bosch et al., 1997; Puget et al., 1999a).

The reported large rearrangements range from 0.5 to 23.8 kb and are spread over the entire *BRCA1* gene. In addition, three of the reported deletions remove portions of the 5′ end of *BRCA1* and its surrounding regions (Swensen et al., 1997; Puget et al., 1999a). Thus, screening *BRCA1* for large rearrangements requires a panoramic view of the 81 kb of the gene and its flanking regions. With this goal, we have chosen the dynamic molecular combing technique associated with fluorescence in situ hybridization (FISH) (Bensimon et al., 1994; Michalet et al., 1997). Combing relies on homogeneous stretching of DNA molecules, which is achieved using a simple apparatus that dips a silanized coverslip into DNA solution and pulls it out at a constant vertical speed to obtain DNA molecules aligned and stretched at a constant rate of 2 kb/μm. We have developed a FISH strategy inspired by Florijn et al. (1995) and Michalet et al. (1997). A complex *BRCA1* full four-color bar code was therefore designed, allowing for analysis of the whole region divided into 10 fragments.

Here we report the usefulness of this approach by analyzing four *BRCA1* rearrangements previously described (Montagna et al., 1999; Puget et al., 1999a, b). In addition, the efficiency of the approach was examined on a series of 10 breast and ovarian cancer cases that tested negative for the *BRCA1/2* point mutation.

MATERIALS AND METHODS

Preparation of Genomic DNA for Combing

Total human genomic DNA was obtained from peripheral blood lymphocytes or lymphoblastoid cell lines. Because use of high-molecular-weight DNA is a limiting factor, cells were embedded in low-melting agarose blocks, as described previously (10^6 cells per block, equivalent to 6 μg of DNA) (Michalet et al., 1997). DNA was diluted to approximately 1.5 μg/mL in 400 mM MES-NaOH at pH 5.5 (Sigma, St. Louis, MO). Combing was done using the Molecular Combing Apparatus™ (Institut Pasteur, Paris, France) and with silanized surfaces (22 × 22 mm) (Michalet et al., 1997).

Labeling of Probes

Two types of probes were used: genomic DNA cloned in P_1-derived artificial chromosomes (PACs) 103O14 and 44B1 (from chromosome 17 library, LLNL, Livermore, CA) (Brown et al., 1995) or cosmid ICRFc105D06121 (from RZPD, Berlin, Germany) (Brown et al., 1996) and products of long-range PCR (LR). PACs and the cosmid were grown and DNA extracted with the very low copy cosmid purification protocol (Qiagen, Valencia, CA). Long-range PCR products were generated from PACs with the Expand™ long template PCR system (Roche Diagnostics, Mannheim, Germany). Primers used are available on the Institut Curie Web site (http://www.curie.net/genetique) or upon request. Each probe (1 μg) was labeled by random priming (BIOPRIME DNA Labeling System; GIBCO, Gaithersburg, MD): PAC 103O14 with fluorescein-11-dUTP (FluoroGreen; Amersham, Piscataway, NJ), cosmid with digoxigenin-11-dUTP (Roche Diagnostics), LR9-12 (6,5 kb) and LR24-3′ (10 kb) with biotin-14-dCTP (GIBCO), and LR16-19 (8 kb) with both biotin and digoxigenin. Each probe was purified with the QIAquick PCR purification kit (Qiagen). Nine hundred nanograms of each probe was ethanol-precipitated with 12× excess human Cot-1 DNA (1 mg/mL; GIBCO).

Hybridization and Probe Detection

Hybridization was performed as previously described (Michalet et al., 1997). Biotinylated, digoxigenin- and fluorescein-labeled probes were detected with AMCA, Texas Red, and fluorescein, respectively. Four or five successive layers of antibodies were used as follows: (1) and (3) AMCA-avidin 1/10 (Vector Laboratories, Burlingame, CA) plus anti-dig–Texas Red (mouse) 1/50 (Jackson Immunoresearch, West Grove, PA) plus anti-fluorescein (rabbit) 1/10 (Molecular Probes, Eugene, OR); (2) and (4) biotinylated anti-avidin 1/50 (Vector Laboratories) plus anti-mouse–dig 1/10 (Roche Diagnostics) plus anti-rabbit–fluorescein (goat) 1/10 (Vector Laboratories); and (5) AMCA-avidin 1/10 plus anti-dig–Texas Red (mouse) 1/50. Washes and antibody incubations were carried out as previously described (Michalet et al., 1997).

Image Analyses

Full signals, corresponding to the *BRCA1* gene, were observed under an epifluorescence Leica DMRB microscope and captured with IPLab Spectrum-SU2 software (Vysis, Downers Grove, IL) using a NU 200 CCD camera (Photometrics, Tucson,

AZ). Image analyses were performed with CartographiX software (X. Michalet, Institut Pasteur), which converts fluorescent spots into intensity curves. Measurements of each fragment in a series of signals are reported on each histogram, resulting in an estimation of the fragment size with its standard deviation due to the constant stretching rate of 2 kb/μm. CartographiX draws normal curves, thereby identifying two putative allele populations.

RESULTS

The *BRCA1* Bar Code Design

The genomic structure of the *BRCA1* region and the probes used are shown in Figure 1a. Counterstaining of the *BRCA1* region was performed with a PAC. A cosmid was used as a 5′ probe and a long-range PCR product, LR24-3′, at the 3′ end. Since we considered the interval between these probes to be too large (76.25 kb) (Smith et al., 1996) to detect rearrangements of less than 5 kb—and to improve their localization—two additional LR products, LR9-12 and LR16-19, were hybridized (Fig. 1a). By using combinations of different fluorescences of probes that divided the *BRCA1* region into 10 fragments, the following full four-color bar code was expected: [NBR1]–cosmid–[NBR2]–cosmid–[intron 2-intron 8]–LR9-12–[intron 12-intron 15]–LR16-19–[intron 19-exon 24]–LR24-3′.

BRCA1 Bar Code of Two DNA Controls

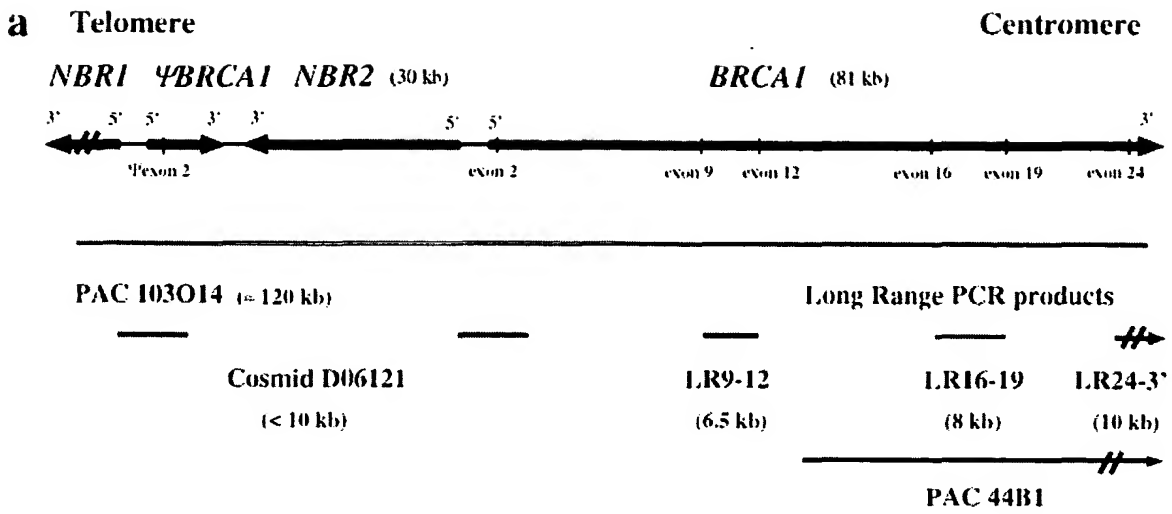
DNA samples from IC918 and IC1063 families, each harboring a point mutation in *BRCA1* or *BRCA2* (Stoppa-Lyonnet, unpublished data), were used as controls to optimize the *BRCA1* bar code strategy. The DNA samples were combed and simultaneously hybridized with the five probes. After screening each slide, full signals of the *BRCA1* gene were compiled and organized with the same orientation, to facilitate qualitative and quantitative analyses. Two IC918 full signals are presented in Figure 1b. Hybridization of the five probes was observed, leading to the expected bar code of the *BRCA1* region.

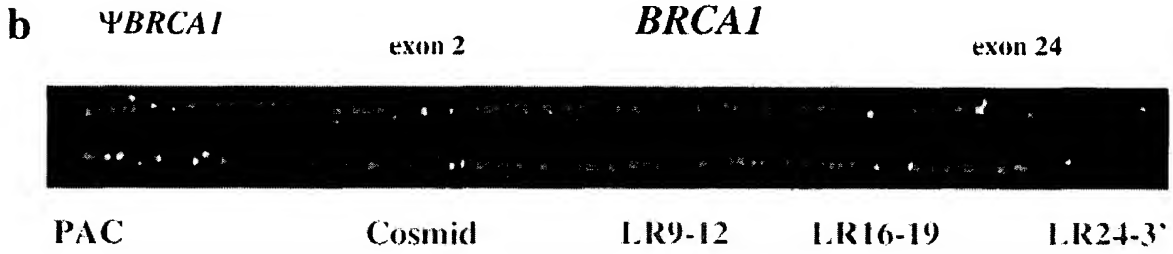
The *BRCA1* gene was divided for measurement into seven fragments ranging from 6.5 to 24 kb. Signals of each fragment were then measured by CartographiX software. The seven IC918 histograms are shown in Figure 1c. A single peak with a normal shape was seen for each histogram. The expected and measured sizes (averages and standard deviations) of the seven fragments are reported in Table 1. The observed sizes of LR signals were close to their expected sizes, with standard deviations ranging from 0.4 to 1.6 kb. For green PAC signals, the observed sizes were lower than the expected sizes deduced from the genomic sequence (Smith et al., 1996). For these larger fragments, standard deviations ranged from 2.0 to 3.0 kb. The variability of a measured fragment is defined as the standard deviation divided by average size. For both controls, the observed variations range from 6% to 18% (average, 13%; Table 1). Variability may be linked to the beaded aspect of signals likely due to the lack of hybridization signal on repeated sequences and to the mild heterogeneity of the coverslip treatment (Michalet et al., 1997). Because the variability is independent of fragment size, however, the smaller the studied fragment, the higher the precision of its measurement (Table 1).

Finally, when studying a rearrangement, the two fragments corresponding to the two allele populations are expected to be detected through the identification of two peaks on a histogram, discriminated by twice the standard deviation (2SD). Considering the 6.5 kb of LR9-12 and the 24 kb of the largest green PAC fragment, [intron 2-intron 8], deletions in these fragments of at least 1.7 kb and 6.2 kb, respectively, should be visible. According to the experimental protocol, 100 equivalent genomes per slide are expected, a quantity sufficient for thorough analysis. Indeed, mean fragment sizes for control measurements are similar for 20, 40, or 70 signals. If the quantity of signals does not appear to be limiting, the quality of DNA is crucial to avoid broken signals.

Detection of a Large Deletion Removing One of the Probes

Puget et al. have identified by Southern blotting a 23.8-kb deletion that removes exons 8 to 13 of *BRCA1* in the IC568 family (Puget et al., 1999a). After combing and hybridization, this large *BRCA1* deletion was detected easily with few fields of view. Some full signals without LR9-12 were observed, thereby allowing us to distinguish between the two alleles (Fig. 2a). To evaluate the size of the deletion better, the region between cosmid and LR16-19, which includes both breakpoints, was measured on the entire series of signals (Fig. 2b; Table 1). As expected, two peaks were obtained, representing the two allele populations. The size and the standard deviation of the deletion were therefore estimated as the difference between the two averages: 20.0 ± 9.6 kb. The known size of 23.8 kb falls within this interval. The large standard


a
Telomere
Centromere
NBR1 ΨBRCA1 NBR2 (30 kb)
BRCA1 (81 kb)
ΨExon 2
exon 2
exon 9
exon 12
exon 16
exon 19
exon 24
PAC 103O14 (≈ 120 kb)
Long Range PCR products
Cosmid D06121
(< 10 kb)
LR9-12
(6.5 kb)
LR16-19
(8 kb)
LR24-3'
(10 kb)
PAC 44B1


b
ΨBRCA1
exon 2
BRCA1
exon 24
PAC
Cosmid
LR9-12
LR16-19
LR24-3'

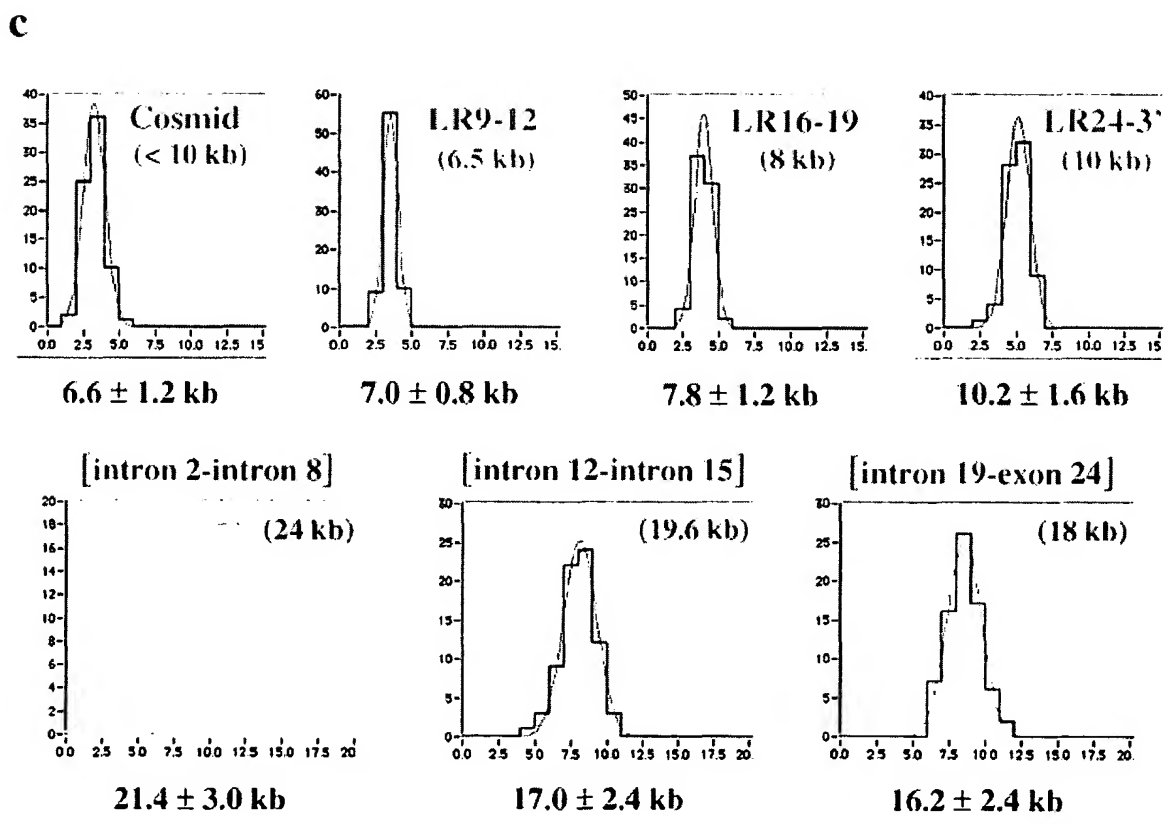

c
Cosmid
(< 10 kb)
6.6 ± 1.2 kb
LR9-12
(6.5 kb)
7.0 ± 0.8 kb
LR16-19
(8 kb)
7.8 ± 1.2 kb
LR24-3'
(10 kb)
10.2 ± 1.6 kb
[intron 2-intron 8]
(24 kb)
21.4 ± 3.0 kb
[intron 12-intron 15]
(19.6 kb)
17.0 ± 2.4 kb
[intron 19-exon 24]
(18 kb)
16.2 ± 2.4 kb

Figure 1

deviation is linked to the size of the analyzed fragment (expected to be 50 kb for the normal allele and 26.4 kb for the deleted allele). From these results, it appears that when a fluorescent fragment is removed, the color bar code strategy gives qualitative information: the existence of large *BRCA1* rearrangements, their locations, and the estimation of their minimal sizes. Nevertheless, it remains difficult to evaluate rearrangement size precisely.

Detection of a 6-kb Duplication

A duplication, recently identified by lymphocyte transcript analysis and precisely characterized by PCR, was shown to be a 6-kb duplication of exon 13 and its surrounding regions (Puget et al., 1999b). One of the DNA samples harboring this duplication, IARC3653, was analyzed. Through microscope screening and by signal alignment, it was difficult to distinguish two allele populations (data not shown). Measurements of the fragment of interest, [intron 12–intron 15], however, yielded two peaks (Fig. 3a; Table 1). The size of the duplication was estimated at 5.8 ± 1.8 kb, compared with the expected 6 kb. This result confirms the resolution capacity of this approach of at least 2SD, which was expected to be 5.1 kb for this 19.6-kb fragment.

Detection of Two 3-kb Deletions

A 3-kb deletion, which removes exon 17, was identified in the Italian family UP-B74 (Montagna et al., 1999). When UP-B74 signals were aligned, it was possible to distinguish several full signals with a shortened LR16-19 probe (Fig. 3b). Measurements of this fragment showed two peaks putatively associated with two allele populations and with average lengths differentiated by 3.2 ± 0.8 kb (Fig. 3c; Table 1). We expected to detect the UP-B74 deletion, since it is located in a fragment of 8 kb with a 2SD of 2.1 kb. This value is in agreement with our estimation of the resolution for the detection of rearrangements that fall within one probe.

To determine whether we are able to detect a rearrangement of a size less than 2SD, we examined the IC788 DNA carrying a 3-kb deletion removing exon 15 within the 19.6-kb fragment [intron 12–intron 15] (Puget et al., 1999a). Measurements of the fragment of interest showed two peaks, leading to the estimation of the size of the deletion at 3.6 kb; however, there was a high standard deviation of 2.8 kb (Fig. 3d; Table 1). Whereas the detection of a 3-kb deletion in an 8-kb fragment is feasible, that of a 3-kb deletion in a 20-kb fragment is at the limit of the resolution of this approach.

Identification of an Unreported *BRCA1* Duplication by Color Bar Code Screening

The color bar code strategy was applied in the analysis of a series of 10 French breast and ovarian cancer families that tested negative for a point mutation in *BRCA1/2*. IC827 DNA showed several full signals larger than the others. Through alignment, the enlarged region appeared to be located between cosmid and LR9-12 (Fig. 4a). Measurements of this region gave two peaks, suggesting the existence of at least a 10-kb duplication in one allele (Fig. 4b; Table 1). To confirm this observation, cosmid and LR9-12 were substituted with two additional probes: LR2-3 and LR5-8 (Fig. 4a). The region between exons 3 and 8 appeared to be duplicated in tandem, as shown by the chimeric

Figure 1. a: The *BRCA1* region and the probes used for bar coding of the region. The *BRCA1* gene is spread over 81 kb (Smith et al., 1996) and has a common promoter with the *NBR2* gene (next to *BRCA1* gene 2), which covers approximately 30 kb (Xu et al., 1997). The *BRCA1* pseudogene (*ΨBRCA1*) lies next to *NBR2* and corresponds to a partial duplication of *BRCA1*, from the promoter region to intron 2 (Barker et al., 1996; Brown et al., 1996). The *NBR1* gene (Next to *BRCA1* gene *1*) is located 5′ to *ΨBRCA1* and may be phylogenetically at the origin of *NBR2* (Campbell et al., 1994; Xu et al., 1997). The figure is not drawn to scale and is adapted from Brown et al. (1996) and Xu et al. (1997). Probes used for the full four-color bar code of *BRCA1*: the PAC 103O14 insert covers the region between the first exons of *NBR1* and the 3′ *BRCA1* UTR (exon 24), which is approximately 120 kb long (Brown et al., 1995, 1996; Smith et al., 1996; Xu et al., 1997). By fluorescence in situ hybridization, this clone hybridized on chromosome 17 and also on chromosome 4, suggesting that this PAC is chimeric. The cosmid ICRFc105D06121 clone used in our experiment is rearranged, since the insert was found to be less than 10 kb (Brown et al., 1996; Xu et al., 1997). By sequencing, one end was localized at the beginning of *BRCA1* intron 2. Considering the size of the insert, the other end was expected to lie in intron 1 of *NBR2*. Thus, our recombinant clone does not spread over the entire *NBR2* gene, as previously described (Brown et al., 1996; Xu et al., 1997). In addition, our clone was expected to cross-hybridize with *ΨBRCA1*. The PAC 44B1 insert has been located in the 3′ *BRCA1* region (Brown et al., 1995). Its *BRCA1* end lies in intron 13, and the other end is expected to be located upstream of the *BRCA1* 3′ end (Smith et al., 1996). Because this PAC appeared to be too large to be used as a 3′ probe, a long-range PCR product, LR24-3′, was generated from this clone, covering *BRCA1* exon 24 to 10 kb downstream. Two other probes, generated by long-range PCR from PAC 103O14, were added: LR9-12, from exons 9 to 12 covering 6.5 kb, and LR16-19, from exons 16 to 19 covering 8 kb. b: IC918 control DNA: two full signals showing the bar code of *BRCA1* (IC: Institut Curie). After screening the IC918 slide, 74 full signals were captured. To facilitate their view and analyses, full signals were aligned. All probes hybridized well. Since the four shortest probes overlap with the counterstaining PAC, color mixes should have been obtained at their locations. The lack of PAC hybridization signal in these overlapping regions is probably due to its low molarity compared with that of the four other probes, since the same quantity was used for each of the five probes. c: The seven IC918 histograms, with microns in abscissa and the number of signals in ordinate. The analyzed fragment and its expected size are reported in the top part of each histogram. The observed sizes and their standard deviations are indicated in kilobases below each histogram.

TABLE 1. Measurements of the Seven Fragments Covering the *BRCA1* Gene*

Hybridization color probes on *BRCA1* Expected fragment sizes	Red cosmid (exons 1–2) < 10 kb	Green [Intron 2–intron 8][a] 24 kb	Blue LR9-12 6.5 kb	Green [Intron 12-intron 15][a] 19.6 kb	Violet LR16-19 8 kb	Green [Intron 19–exon 24][a] 18 kb	Blue LR24-3' 10.1 kb	Full signal number
IC918 (control 1)	6.6 ± 1.2 kb (18.2%)	21.4 ± 3.0 kb (14.0%)	7.0 ± 0.8 kb (11.4%)	17.0 ± 2.4 kb (14.1%)	7.8 ± 1.2 kb (15.4%)	16.2 ± 2.4 kb (14.8%)	10.2 ± 1.6 kb (15.7%)	74 (14.8%)
IC1063 (control 2)	5.6 ± 1.0 kb (17.8%)	19.2 ± 2.2 kb (11.4%)	6.2 ± 0.4 kb (6.4%)	15.0 ± 2.0 kb (13.3%)	7.4 ± 0.8 kb (10.8%)	14.2 ± 2.0 kb (14.0%)	9.4 ± 0.8 kb (8.5%)	47 (11.7%)
IC568	5.4 ± 1.2 kb (22.2%)		Two allele populations		7.4 ± 0.6 kb (8.1%)	14.8 ± 2.6 kb (17.6%)	9.6 ± 0.8 kb (8.3%)	62 (14.0%)
(23.8-kb deletion, from exons 8 to 13)			32 deleted alleles 24.8 ± 4.0 kb (16.1%) 30 normal alleles 44.8 ± 5.6 kb (12.5%) del = 20.0 ± 9.6 kb					1st peak 2nd peak
IARC3653	6.0 ± 1.2 kb (20.0%)	21.4 ± 2.2 kb (10.3%)	6.6 ± 0.4 kb (6.1%)	19.8 ± 3.4 kb (17.2%)	8.0 ± 1.0 kb (12.5%)	15.6 ± 2.2 kb (14.1%)	9.6 ± 1.4 kb (14.6%)	54 (13.5%)
(6-kb duplication, exon 13)				16.8 ± 0.8 kb (4.8%) 22.6 ± 1.0 kb (4.4%) dup = 5.8 ± 1.8 kb				1st peak: 27 2nd peak: 27
UP-B74	6.0 ± 1.4 kb (23.3%)	19.8 ± 2.8 kb (14.1%)	6.4 ± 0.4 kb (6.2%)	16.0 ± 2.2 kb (13.7%)	6.4 ± 1.6 kb (25.0%)	15.0 ± 2.8 kb (18.7%)	9.6 ± 0.8 kb (8.3%)	46 (15.6%)
(3-kb deletion, exon 17)					4.8 ± 0.4 kb (8.3%) 8.0 ± 0.4 kb (5.0%) del = 3.2 ± 0.8 kb			1st peak: 21 2nd peak: 25
IC788	5.2 ± 1.2 kb (23.0%)	17.4 ± 2.6 kb (14.9%)	6.2 ± 0.4 kb (6.5%)	14.2 ± 2.4 kb (16.9%)	7.6 ± 0.4 kb (5.3%)	13.2 ± 1.8 kb (13.6%)	9.6 ± 0.4 kb (4.2%)	46 (12.1%)
(3-kb deletion, exon 15)				12.6 ± 0.8 kb (6.3%) 16.2 ± 2.0 kb (12.3%) del = 3.6 ± 2.8 kb				1st peak: 20 2nd peak: 26
IC827	5.6 ± 1.2 kb (21.4%)	24.6 ± 6.6 kb (26.8%)	6.4 ± 0.4 kb (6.2%)	15.2 ± 2.8 kb (18.4%)	7.4 ± 0.8 kb (10.8%)	14.8 ± 2.4 kb (16.2%)	9.6 ± 0.8 kb (8.3%)	57 (15.4%)
(unreported duplication)		19.2 ± 2.2 kb (11.4%) 31.0 ± 3.8 kb (12.2%) dup = 11.8 ± 6.0 kb						1st peak: 31 2nd peak: 26

*For each patient, the fragment including the rearrangement is detailed. Except for IC568, where two allele populations are distinguishable, the first line indicates the average size of signals. The second and third lines indicate the mean size of the two fragment populations. Percentages reflect the variability observed for each fragment and for the full signal in the last column.
[a]PAC hybridization fragments (Fig. 1).

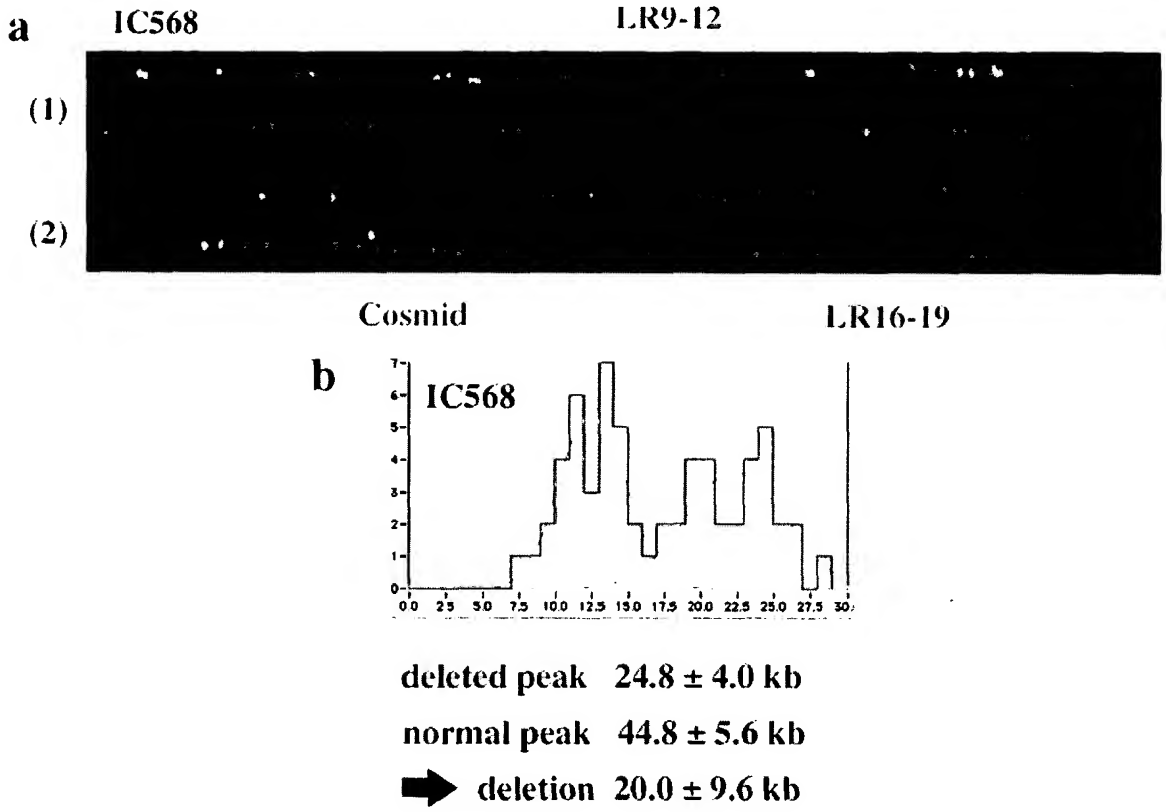


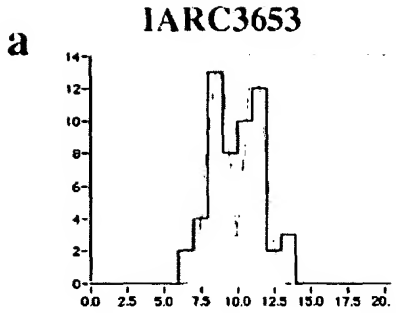
Figure 2. IC568 DNA harboring a 23.8-kb deletion (from exons 8 to 13) (Puget et al., 1999a). a: Two normal full signals (1) and two full signals without LR9-12 probe (2). Sixty-two full signals were obtained—30 signals with LR9-12 and 32 without—allowing the distinction between the two allele populations and yielding qualitative information about the nature of the rearrangement, its localization, and minimal size. b: Histogram of measurements between cosmid and LR16-19 on both normal and deleted signals. Two peaks representing the two allele populations are detectable. The size of the deletion is obtained by determining the difference between the two peaks.

hybridization signal of LR5-8 and a small part of LR2-3. Fine characterization using long-range PCR in the region of interest showed a 17,247 bp duplication of exons 3 to 8 (nucleotides 11967–29213 [GenBank accession number L78833]), which is the result of a homologous recombination between an Sq Alu in intron 2 and an Sp Alu sequence in intron 8, which share 83% homology (data not shown). Under the microscope, the other DNA samples did not seem to have large rearrangements, and signal measurements did not detect additional rearrangements (data not shown).

DISCUSSION

We have shown that the strategy combining DNA combing and full four-color bar coding is useful for detecting large *BRCA1* rearrangements, since a panoramic view of the gene and its surrounding regions can be obtained. The detection of rearrangements as small as 2 to 6 kb with respect to the normal size of the studied fragment is achieved when the *BRCA1* region is divided into 10 fragments. The efficiency of our approach was verified with the detection of a previously unreported *BRCA1* duplication when screening 10 breast and ovarian cancer cases that had tested negative for the *BRCA1/2* point mutation. This bar code is expected to be able to detect 80% of the large rearrangements previously reported in the literature. Hybridization with additional probes and use of other colors, splitting the *BRCA1* gene into fragments shorter than 10 kb, should lead to the detection of rearrangements smaller than 2 kb. An optimized bar code and the analysis of a larger series would allow for accurate estimation of the contribution of large rearrangements in the *BRCA1* mutation spectrum.

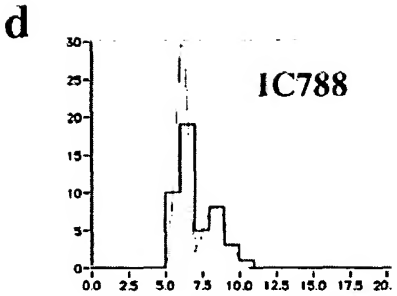
Detection of large rearrangements represents a recurrent challenge in molecular pathology. Southern blotting, which was the first tool used to detect large rearrangements, relies mainly on the detection of abnormal restriction endonuclease frag-



normal peak 16.8 ± 0.8 kb

duplicated peak 22.6 ± 1.0 kb

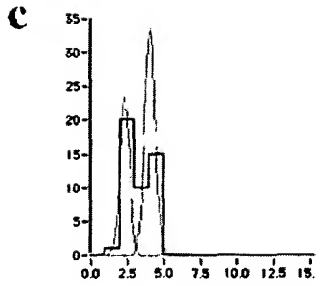
➡ duplication 5.8 ± 1.8 kb



deleted peak 12.6 ± 0.8 kb

normal peak 16.2 ± 2.0 kb

➡ deletion 3.6 ± 2.8 kb



deleted peak 4.8 ± 0.4 kb

normal peak 8.0 ± 0.4 kb

➡ deletion 3.2 ± 0.8 kb

Figure 3. a: IARC3653 DNA harboring a 6-kb duplication (exon 13; IARC: International Agency for Research on Cancer) (Puget et al., 1999b). Histogram of measurements between LR9-12 and LR16-19, [intron 12–intron 15], on 54 full signals. Two peaks are detected. b, c: UP-B74 DNA harboring a 3-kb deletion (exon 17; UP: University of Padova) (Montagna et al., 1999). b: Two normal full signals (1) and two full signals with a smaller LR16-19 probe (2). c: Histogram of LR16-19 measurements on 46 full signals. Two peaks are observed. d: IC788 DNA harboring a 3-kb deletion (exon 15; Puget et al., 1999a). Histogram of [intron 12–intron 15] measurements on 46 full signals. Two peaks are detected.

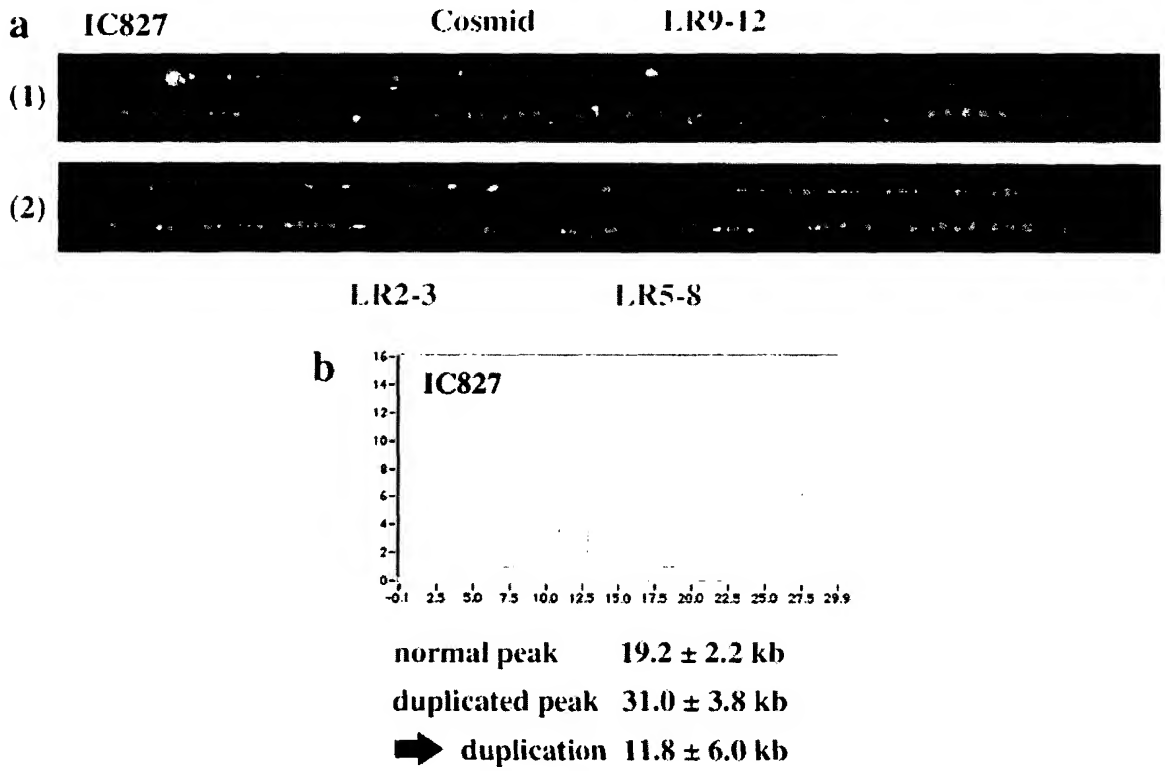


Figure 4. IC827 DNA harboring an unreported duplication. a: IC827 full signals. (1) The lower signal appeared larger than the upper signal for the region between cosmid and LR9-12: [intron 2–intron 8]. (2) Hybridization of two additional long-range products covering exons 2 to 3 of 8.5 kb (LR2-3) and exons 5 to 8 of 7 kb (LR5-8), in substituting for cosmid and LR9-12, respectively. The lower signal contains an additional chimeric hybridization pattern of probes LR5-8/LR2-3, indicating the duplication junction. b: Histogram of [intron 2–intron 8] measurements on 57 full signals. Two peaks representing the two allele populations are detectable.

ments induced by rearrangements. This strategy may fail when the target sequence recognized by the probe also is deleted. The difficulty encountered in distinguishing between the presence of one or two copies in the case of a deletion is augmented when distinguishing between two or three copies in the case of a duplication. A second tool is the analysis of transcript size in the available tissue (lymphocytes in most cases). It requires conditions that are not necessarily fulfilled, however: transcription of the studied gene and stability of altered allele transcripts. The approach described in this study may avoid these disadvantages. In addition, because it is a morphologic approach, the direct observation of complex and likely underreported rearrangements, such as inversions and insertions, becomes possible. This new strategy, until now primarily used in research laboratories, can be implemented for the molecular diagnosis of other genetic diseases as well. Indeed, the bar code of other genes may be designed easily using artificial clones and/or genomic sequence information. Development of software allowing, on the one hand, the automatic capture of signals with the help of motorized microscope stages and, on the other hand, their automatic measurements would favor the use of this new strategy by clinical genetics laboratories.

ACKNOWLEDGMENTS

We are grateful to Xavier Michalet and Pierre Stanislavski (Institut Pasteur) for their introduction to CartographiX software. We thank Emma D'Andrea (University of Padova) for critical reading and Julian Lange for linguistic revision. We are indebted to the families for participating in this study. S.G. is supported by a fellowship from MENRT. A.A. thanks the Ligue Nationale Contre le Cancer, the Comité National et Comité de l'Oise, and the FEGEFLUC. N.P. is a fellow of the Ligue Nationale Contre le Cancer de Haute-Savoie. M.M. is a member of the Italian Hereditary Breast-Ovarian Cancer Consortium/AIRC-Coordinated Project and is supported by a fellowship from

the Fondazione Cassa di Risparmio di Padova e Rovigo.

REFERENCES

Barker DF, Liu X, Almeida ERA. 1996. The *BRCA1* and *IA13B* promoters are parallel elements of a genomic duplication at 17q21. Genomics 38:215–222.

Bensimon A, Simon A, Chiffaudel A, Croquette V, Heslot F, Bensimon D. 1994. Alignment and sensitive detection of DNA by a moving interface. Science 265:2096–2098.

Brown MA, Jones KA , Nicolai H, Bonjardim M, Black D, McFarlane R, De Jong P, Quirk JP, Lehrach H, Solomon E. 1995. Physical mapping, cloning, and identification of genes within a 500 kb region containing *BRCA1*. Proc Natl Acad Sci U S A 92:4362–4366.

Brown MA, Xu CF, Nicolai H, Griffiths B, Chambers JA, Black D, Solomon E. 1996. The 5′ end of the *BRCA1* gene lies within a duplicated region of human chromosome 17q21. Oncogene 12: 2507–2513.

Campbell IG, Nicolai HM , Foulkes WD, Senger G, Stamp GW, Allan G, Boyer C, Jones K, Bast Jr RC, Solomon E, Trowsdale J, Black DM. 1994. A novel gene encoding a B-box protein within the *BRCA1* region at 17q21.1. Hum Mol Genet 3:589–594.

Carson N, Gilpin C, Hunter A, Allanson J, Aubry H. 1999. An in-frame deletion of *BRCA1* exon 20 in a family with early-onset breast cancer and ovarian cancer. Am J Hum Genet 65:A1610.

Florijn RJ, Bonden LAJ, Vrolijk H, Wiegant J, Vaandrager JW, Baas F, den Dunnen JT, Tanke HJ, van Ommen GJB, Raap AK. 1995. High-resolution DNA Fiber-FISH for genomic DNA mapping and colour bar-coding of large genes. Hum Mol Genet 4: 831–836.

Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT, Weber B, Lenoir G, Chang-Claude J, Sobol H, Teare MD, Struewing J, Arason A, Scherneck S, Peto J, Rebbeck TR, Tonin P, Neuhausen S, Barkardottir R, Eyfjord J, Lynch H, Ponder BAJ, Gayther SA, Birch JM, Lindblom A, Stoppa-Lyonnet D, Bignon Y, Borg A, Hamann U, Haites N, Scott RJ, Maugard CM, Vasen H, Seitz S, Cannon-Albright LA, Schofield A, Zelada-Hedman M, the Breast Cancer Linkage Consortium. 1998. Genetic heterogeneity and penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. Am J Hum Genet 62:676–689.

Gayther SA, Ponder BA. 1997. Mutations of the *BRCA1* and *BRCA2* genes and the possibilities for predictive testing. Mol Med Today 3:168–174.

Michalet X, Ekong R, Fougerousse F, Rousseaux S, Schurra C, Hornigold N, van Slegtenhorst M, Wolfe J, Povey S, Beckmann JS, Bensimon A. 1997. Dynamic molecular combing: stretching the whole human genome for high-resolution studies. Science 277:1518–1523.

Montagna M, Santancatterina M, Torri A, Menin C, Zullato D, Chieco-Bianchi L, D'Andrea E. 1999. Identification of a 3 kb Alu-mediated *BRCA1* gene rearrangement in two breast/ovarian cancer families. Oncogene 18:4160–4165.

Payne SR, King M-C. 1997. Complex *BRCA1* germline deletion and rearrangement of 1028 bp leads to deletion of exon 3 and disease predisposition: splicing by exon definition in *BRCA1*. Am J Hum Genet 61:A425.

Petrij-Bosch A, Peelen T, van Vliet M, van Eijk R, Olmer R, Drüsedau M, Hogervost FBL, Hageman S, Arts PJW, Ligtenberg MJL, Meijers-Heijboer H, Klijn JGM, Vasen HFA, Cornelisse CJ, van't Veer LJ, Bakker E, van Ommen GJB, Devilee P. 1997. *BRCA1* genomic deletions are major founder mutations in Dutch breast cancer patients. Nat Genet 17:341–345.

Puget N, Torchard D, Serova-Sinilnikova O, Lynch HT, Feunteun J, Lenoir GM, Mazoyer S. 1997. A 1 kb Alu-mediated germ-line deletion removing *BRCA1* exon 17. Cancer Res 57:828–831.

Puget N, Stoppa-Lyonnet D, Sinilnikova OM, Pages S, Lynch HT, Lenoir GM, Mazoyer S. 1999a. Screening for germ-line rearrangements and regulatory mutations in *BRCA1* led to the identification of four new deletions. Cancer Res 59:455–461.

Puget N, Sinilnikova OM, Stoppa-Lyonnet D, Audoynaud C, Pages S, Lynch HT, Goldgar D, Lenoir GM, Mazoyer S. 1999b. An Alu-mediated 6 kb duplication in the *BRCA1* gene: a new founder mutation? Am J Hum Genet 64:300–302.

Rohlfs EM, Yang Q, Skrzynia C, Graham ML, Silverman LM. 1999. Complex rearrangement in *BRCA1* results in the in-frame loss of exons in the BRCT domain. Am J Hum Genet 65:A1792.

Rohlfs EM, Puget N, Graham ML, Weber BL, Garber JE, Skrzynia C, Halperin JL, Lenoir GM, Silverman LM, Mazoyer S. 2000. An Alu-mediated 7.1 kb deletion of *BRCA1* exons 8 and 9 in breast and ovarian cancer families that results in alternative splicing in exon 10. Genes Chromosomes Cancer 28:300–307.

Smith TM, Lee MK, Szabo CI, Jerome N, McEuen M, Taylor M, Hood L, King MC. 1996. Complete genomic sequence and analysis of 117 kb of human DNA containing the gene *BRCA1*. Genome Res 6:1029–1049.

Stoppa-Lyonnet D, Puig PL, Essioux L, Pages S, Ithier G, Ligot L, Fourquet A, Salmon RJ, Clough KB, Pouillard P, the Institut Curie Breast Cancer Group, Bonaiti-Pellie C, Thomas G. 1997. *BRCA1* sequence variations in 160 individuals referred to a breast/ovarian family cancer clinic. Am J Hum Genet 60:1021–1030.

Swensen J, Hoffman M, Skolnick MH, Neuhausen SL. 1997. Identification of a 14 kb deletion involving the promoter region of *BRCA1* in a breast cancer family. Hum Mol Genet 6: 1513–1517.

Xu CF, Brown MA, Nicolai H, Chambers JA, Griffiths B, Solomon E. 1997. Isolation and characterization of the *NBR2* gene which lies head to head with the human *BRCA1* gene. Hum Mol Genet 6:1057–1062.